AFRL-RH-WP-TR-2016-0088

ANALYSIS OF DISPLAY LATENCY FOR 3D PERCEPTUAL EXPERIMENTS

ELIZABETH FOX and LESLIE M. BLAHA

711 HPW/RHCV

NOVEMBER 2016

Interim Report

STINFO COPY

AIR FORCE RESEARCH LABORATORY
711TH HUMAN PERFORMANCE WING,
AIRMAN SYSTEMS DIRECTORATE,
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433
AIR FORCE MATERIEL COMMAND
UNITED STATES AIR FORCE

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report was cleared for public release by the 88th Air Base Wing Public Affairs Office and is available to the general public, including foreign nationals. Copies may be obtained from the Defense Technical Information Center (DTIC) (https://www.dtic.mil).

AFRL-RH-WP-TR-2016-0088 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//signed//
ROY R. RASMUSSEN
Program Manager
Battlespace Visualization Branch

//signed//
ROBERT C. MCKINLEY
Chief, Battlespace Visualization Branch
Warfighter Interface Division

//signed//
WILLAIM E. RUSSELL
Chief, Warfighter Interface Division
Airman Systems Directorate

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
23-11-16	Interim	1 May 2014 – 1 May 2016

4. TITLE AND SUBTITLE

ANALYSIS OF DISPLAY LATENCY FOR 3D PERCEPTUAL EXPERIMENTS

5a. CONTRACT NUMBER
FA8650-14-D-6501/0005

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
61102F

6. AUTHOR(S)
Elizabeth Fox
Leslie M. Blaha

5d. PROJECT NUMBER
2313

5e. TASK NUMBER
2313CV

5f. WORK UNIT NUMBER
H0JW (2313CV02)

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
711 HPW/RHCV
2255 H Street, Bldg 248
Wright-Patterson AFB, OH 45433

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Materiel Command
Air Force Research Laboratory
711th Human Performance Wing
Airman Systems Directorate
Warfighter Interface Division
Battlespace Visualization Branch
Wright-Patterson Air Force Base, OH 45433

10. SPONSORING/MONITORING AGENCY ACRONYM(S)
711 HPW/RHCV

11. SPONSORING/MONITORING AGENCY REPORT NUMBER(S)
AFRL-RH-WP-TR-2016-0088

12. DISTRIBUTION/AVAILABILITY STATEMENT
DISTRIBUTION STATEMENT A. Approved for public release: distribution unlimited.

13. SUPPLEMENTARY NOTES
88ABW Cleared 12/08/2016; 88ABW-2016-6354. Report contains color.

14. ABSTRACT
The purpose of this invesgitation was to test if timing variability in a 3D stereo display was correlated with the displayed degree of stereo depth cue. A series of recordings of display timing were made under various levels of experimental manipulation to test for correlation between dimensions. Small correlations were identified in the device, as well as systematic delays in recorded timing over scheduled timing. No meaningful correlation was found between display timing and human response times. We conclude that any perception of a correlation is due to the human perceptual system and should not prevent further use of the tested 3D display system for further empricial research in 711 HPW/RHCV.

15. SUBJECT TERMS
Display Technology; Human Factors; Human Perception, Stereoscopic display; Stereoscopic perception

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT:	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON (Monitor)
a. REPORT	b. ABSTRACT	c. THIS PAGE			Roy R. Rasmussen
Unclassified	Unclassified	Unclassified	SAR	22	19b. TELEPHONE NUMBER *(Include Area Code)*

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18

Table of Contents

List of Figures

Figure **Page**

List of Tables

Table **Page**

1. Introduction

During human data collection for a 3D perceptual task, several participants self-reported the experience that objects with a greater degree of depth took longer to appear on the screen. That is, they perceived inter-stimulus interval (ISI), the time interval from the offset of the fixation cross to the onset of the stimulus, as possibly dependent on the amount of depth at which the stimulus was presented (see Figure 1 for the trial event sequence). Specifically, these participants reported that a long ISI seemed to indicate a high degree of depth in the following stimulus while a short ISI could have been evidence that the stimulus to follow would have little-to-no stereo disparity. Their reports suggest a positive correlation between the amount of stereo disparity in the target stimulus and the amount of time it takes to display the image on the screen.

The existence of such a correlation in experimental timing would pose a problem for human behavioral experimentation because the presence of such a correlation might bias the perceptual decision making processes of the observers. That is, if the participants are given an ISI timing cue before the stimulus is presented, we are no longer measuring only the response time from the onset of the stimulus or the processing of the 3D/2D stimulus. Rather, their responses reflect some biased combination of the stimulus perception and the perceived timing of the ISI.

The reported apparent correlation between depth and ISI length could be the result of a perceptual phenomenon or a hardware issue within the 3D display. When a stimulus is presented in depth, it will cause a physical verging of the eyes. With a larger disparity cues, more convergence is necessary in order to fuse the stereo images into a single 3D percept. It is possible that the larger physical movements may cause a longer perceived ISI for higher disparity levels. Another possible explanation may be that the hardware used for the stimulus presentation requires more time to render the stimuli that have a larger disparity than those with a low or nonexistent (2D) disparity.

The goal of this effort is to assess if there are timing differences inherent to the 3D stereo display that cause the apparent correlations reported by participants. First, we review the human subjects' experiment during which these reports were made. Then we report a series of recordings and analyses taken on the 3D stereo display to assess the presence of a correlation between display time and stereo disparity level. The results of this work are critical to determining if available 3D displays can be used for human perception research or if they introduce artifacts that might bias performance and our inferences about depth perception. It is critical to have hardware in experiments that do not introduce external biases into our experimental procedures.

In the human perception studies of interest, observers were asked to judge if a single circular object on the screen was presented in depth or not (binary Yes/No decision). Objects appeared around the center of the screen, and responses were made with a two-button mouse. The objects could appear at 20 different stereo depth levels ranging from flat on the screen (0 arcmin disparity) to 22.56 arcmin of disparity popping out toward the observer. The top row in Figure 1 depicts a trial that displays a 2D stimulus. Participants reported a short ISI for these trial types. The bottom row of Figure 1 illustrates a trial when a 3D stimulus was presented (image as it appears on the screen without the stereo lenses). Participants reported a longer ISI when 3D stimuli were presented. Specifically, participants speculated that as the stereo disparity increased, the ISI increased. The level of stereo disparity varied randomly on each trial (0-22.56 arcmin). Participants wore passive polarized 3D lenses for the duration of the experiment.

In the screenshot in Figure 1, lower row, the stereo cue is shown as two images; the amount of offset between the images determines the amount of stereo disparity. As the amount of stereo depth

increases, the two circles separate from one another. To fuse the images into a single percept, participants have to converge their eyes; more change in vergence between the 0 arcmin disparity fixation and the 3D stimulus is needed as the stereo depth cue (distance between the two circles) increases.

Note that because the polarized lenses are passive, there is no possibility of a lens-based hardware asynchrony issue, such as might occur with the use of active polarized lenses. The ISI was set to vary uniformly between 50-750 ms, and duration samples were independent of the stereo disparity level on each trial. This raises additional suspicion that a hardware issue was present. For participants to notice a hardware issue in which the ISI was dependent on trial depth, the hardware would have to systematically add latency correlated with the stereo disparity level to the scheduled random ISI duration.

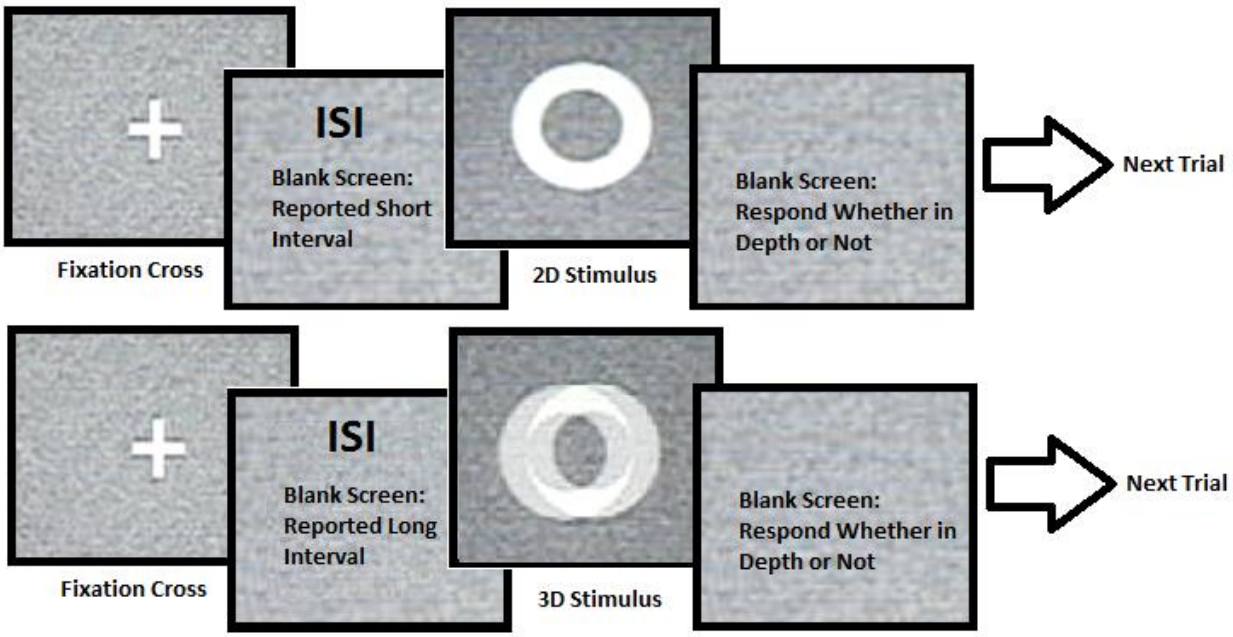

Figure 1. Example trials for both 2D (top row) and 3D (bottom row) stimulus presentation. The screenshots show photos of the display as it would appear without the use of stereo lenses.

In a series of measurement studies, we addressed the potential relationship between the amount of stimulus disparity and the duration of the ISI. The general process was to use high-speed video recording to determine the actual recorded time between the offset of the fixation cross and the onset of the stimulus for every trial. We systematically manipulated the scheduled duration of the ISI and the presentation order of disparity level. Each recording in the series was designed to test possible sources of timing discrepancies.

1. First in Recording 1, we fixed the ISI duration to 50 ms on all trials and increased disparity by one increment (~0.60 arcmin) every 20 trials. This was to provide a baseline for ISI duration without experiment-like variation in the stimulus presentation.
2. In Recording 2, the trial order was randomized to determine if a variable changes of disparity between trials resulted in larger ISI deviations from the programmed, non-random duration (50 ms).
3. Next, in Recording 3, we randomized the ISI duration (sampled from a uniform random variable between 50-750 ms) to detect any systematic deviations from the coded ISI (recorded value between 50-750 ms) due to random changes in stimulus disparity level.
4. Finally, in Recording 4, we fixed the ISI while logging the computer recorded ISI in addition to calculating a camera recorded ISI. The purpose of Recording 4 was to see if deviations of software-recorded ISI and camera-recorded ISI increase as stimulus disparity increases.

To foreshadow our results, the above series of recordings produced some mixed results, including evidence of potential correlations in the machine timing characteristics. Consequently, in Recording 5, we recorded display timing while we collected data from two human observers. If there was no correlation between the human performance and the variability within the recorded display ISI, we could determine the variability found (even if statistically significant) was not practically significant. Hardware variability may yield statistical significance under precise timing measurements but may not exhibit systematic influence on participant decision-making performance. If this is the case, we could conclude the significance of the hardware variability was not large enough for us to change experimental practices. We could deem the hardware sufficient for human experimentation with various 3D stimuli. Therefore, we could continue with data collection as scheduled.

We conducted this investigation only with the 3D stereo display used for the Battlespace Visualization Branch's (711 HPW/RHCV) human studies. Consequently, the exact numerical results herein are specific to that display. However, the methodology we employed may be utilized to test timing properties of other stereoscopic displays of interest.

2. Technical Specifications

2.1. 3D Stereoscopic Display

Stereoscopic images were displayed on an LG 55 in. Class (54.6 in. diagonal) cinema 3D 1080p 240 Hz LED TV for stimulus display. The 3D display utilized circular polarized (film-type patterned retarder; FPR) using a horizontal split type to produce stereoscopic images. Human stereo perception required passive polarized lenses to be worn by the observers.

2.2. Recording Camera

Video recording was done with a High Speed CASIO EX-FH100 Digital Camera. The video and picture files were stored with a Kingston Technology 16GB Micro SD Adapter.

2.3. Software

Video recordings were processed in MATLAB using the Image Processing Toolbox. Data analysis was conducted using R language for statistical computing (R Core Team, 2015). Custom experimental software was programmed in OpenGL for 3D object modeling and image rendering.

3. General Recording Methodology

All recordings were filmed at 240 frames per second (fps). A pause in recording was scheduled approximately every 6 minutes (100 trials) to check spare camera memory capacity. If memory was full, the video recordings were transferred to a laptop computer and erased from the camera device. The camera was positioned at eye-level 120 in. from the display, the same distance as a human participant's chair. A tripod was used to hold the camera in place for all recordings. Grid lines on the camera's user display helped to maintain an accurate center for the recording, critical for the video processing in MATLAB.

In Recording 5, the camera was approximately 120 in. from the center of the display and positioned to the left of the participant, angled to the center of the display. This allowed the participant to be viewing the center of the screen without interference of the camera. The grid lines on the camera display were again used to for center alignment. The stimuli were still clearly visible for all video processing of Recording 5.

4. Data Analysis

All recordings were uploaded to MATLAB. Using a frame-by-frame corner shape detection, we were able to flag the frame number on which the fixation offset occurred. Similarly, the frame containing the onset of the stimulus was determined by circle shape detection. We calculated the number of frames recorded for the ISI by subtracting the frame number of the offset of the fixation from the onset of the stimulus. This process was repeated using MATLAB for each trial within each video recording. We then converted the number of frames to milliseconds (ms). Given the camera records at 240 fps (240 frames per 1000 ms), we could divide ISI frames by 240 to return each ISI in terms of milliseconds. For example, 12 frames = 12/240*1000 = 50 ms.

Throughout the analyses, the key dependent measure of interest is the ISI deviation. We define this as the difference between the scheduled ISI duration and the recorded ISI duration.

Each recording included two planned analyses. First, a t-test was used to compare the recorded ISI durations to the scheduled ISI durations. For Recordings 1-2, where the ISI was fixed to 50 ms, the null hypothesis was that ISI deviation = 0, or the recorded and scheduled ISI were equal. These were one-sample, right-tailed t-tests. For Recordings 3-4 wherein the ISI was variable, paired two-sample, right-tailed t-tests were used.

Multiple regression models were fit to examine potential correlation in the ISI deviations and various experimental conditions. The dependent variable was ISI deviation. Independent variables are specified within each Recording's results sections.

For all analyses, statistical significance was determined against Type I error rate α=.05.

5. Recordings

5.1 Recording 1: Fixed ISI, Fixed Order

The first recording was designed to capture the natural variability in the display timing. A stimulus can only be displayed at times when the screen refreshes. Because trial sequences like those in Figure 1 are dependent on human observer inputs, which may not occur on screen frame flips, the timing of the display will have some inherent variability due to the interaction of observer inputs and the screen refresh timing.

To capture this, all variability in the experimental code was removed. The goal was to identify if there is a systematic difference between the scheduled ISI and the recorded ISI over the course of the experiment. The scheduled ISI was fixed to 50 ms on all trials. Stimulus disparity increased by one increment (~0.60 arcmin) every 20 trials. The lowest disparity level was displayed the first 20 trials (0 arcmin, or the 2D stimulus) and the largest disparity level was displayed on the last 20 trials (22.56 arcmin). A total of thirty-one different disparities were presented for a total of 620 trials.

Using the MATLAB analysis described above (Section 4), we determined the recorded ISI in milliseconds. Using a one-sample, right-tailed t-test, we tested the null hypothesis of no difference between the two ISI durations against the alternative hypothesis that the recorded ISI was longer than the scheduled ISI. The recorded ISI ($M_{Recorded}$=58.20 ms) was significantly greater than the scheduled ISI ($M_{Scheduled}$=50.0 ms), t(619)=18.50, p<.001. Looking at the recorded ISI values (y-axis in Figure 2), although time is continuous, frame refresh rates are not. For the 240 Hz monitor, the screen refreshes every 4.167 ms. The scheduled 50 ms ISI is equal to 12 frames. But as subsequent trials are also dependent on the timing of the human response to the stimulus, the stimulus onset occurs at approximately 12 frames. As Figure 2 illustrates, many of the stimuli appear between 11 and 16 frames. Due to this variability, the hypothesis of interest is not necessarily the degree to which the difference between the scheduled and recorded ISI durations is equal to 0, but if there is a correlation between the recorded ISI and other aspects of the experimental design. That is, does the variability of this timing change over the amount of disparity in the stimuli?

Figure 2 plots disparity against the difference between the scheduled and recorded ISI. Visual inspection indicates a possible increasing trend in the data. Linear regression analysis, shown as the red line in Figure 2, confirmed that disparity level significantly predicted the difference between the scheduled and recorded ISI, F(1, 618)=4.40, p=.036. In the regression model, $B_{Disparity}$=0.03 with an R^2=0.007. So although disparity was a statistically significant predictor of the ISI difference, disparity explained very little of the total variance in the data.

The data from Recording 1 confirm that there is an inherent variability in the stimulus display times, such that the recorded ISI were approximately 50 ms. The positive regression slope also suggests a correlation between the amount of disparity and the amount of variance in this timing. Because stimulus disparity strictly increased over the recording, it is not clear if the increase in timing variance is influenced by the amount of disparity only or is a consequence of time in the experiment. This can be tested by varying the order in which the different disparity levels occur.

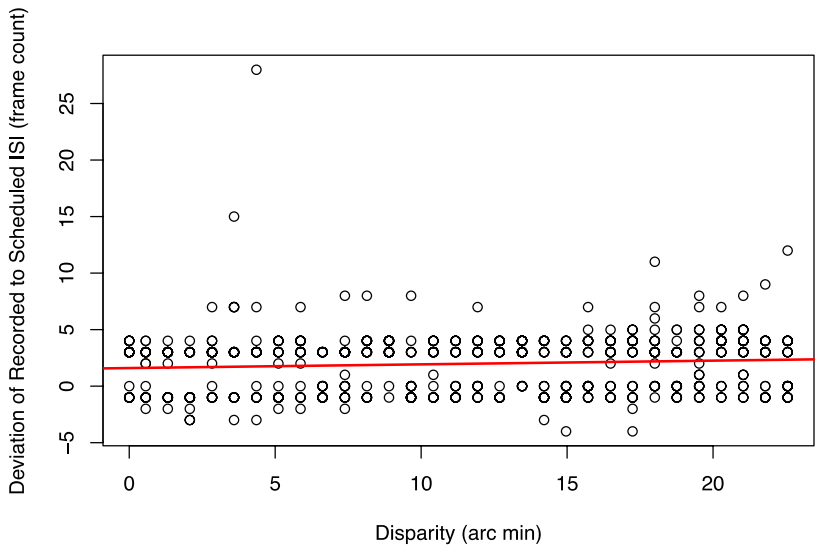

Figure 2. Recording 1: Linear regression between the ISI deviation and stimulus disparity level. Each computed ISI is given a circle icon. The red line gives the regression line.

5.2 Recording 2: Fixed ISI, Random Ordering

Recording 2 added the random trial ordering back into the process but continued to fix the scheduled ISI duration. This allows us to examine the ISI durations as a function of both disparity and experiment duration without the two independent variables being correlated. Each disparity was presented 20 times ranging from 0-22.56 arcmin, same as Recording 1. Now, the stimulus presentation order was randomly shuffled. All other parameters were the same as Recording 1. If the order of the stimuli does not influence the display timing, then the difference between the program's ISI and the recorded ISI should show no correlation.

We again check if the recorded ISI was greater than the scheduled ISI. A one-sample, right-tailed t-test showed that the recorded ISI ($M_{Recorded}$=56.9 ms) was significantly greater than the scheduled ISI ($M_{Scheduled}$=50.0 ms), t(619)=17.67, p<.001.

We analyzed the ISI deviation as a function of both the stimulus disparity and the change in disparity between two subsequent trials. Under the null hypothesis that there is no difference in these values over any other experimental factors, then the difference values should be centered around zero, reflecting the natural variation in frames observed in Recording 1. The difference in ISI values is plotted as a function of the stimulus disparity in Figure 3A, and as a function of the change in disparity between trials in Figure 3B. Data are plotted in frame count.

A multiple regression model explained R^2=0.01 of the variability in the data, F(2, 617)=3.46, p=.03. Stimulus disparity was a significant predictor of ISI deviation, $B_{Disparity}$=0.04, t(617)=2.39, p=0.02. The change in disparity between trials was not a significant predictor of ISI deviation, $B_{DisparityChange}$ =-0.002, t(618)=-0.15, p=0.88.

A linear regression was fit to the ISI deviation as a function of trial order, to parse apart the dependence of camera frame counts of disparity level and experiment duration found in Recording 1. The model showed a marginally significant trend with a nearly zero coefficient value, illustrated in Figure 3C, that experimental duration influenced the difference in ISI values, B_{Order}=0.0009, accounting for R^2=0.005 of the variability, F(1,618)=2.95, p=0.09. Collectively, these results do suggest that the recorded ISI in frame counts is statistically dependent on disparity level of the stimulus rather than experimental duration.

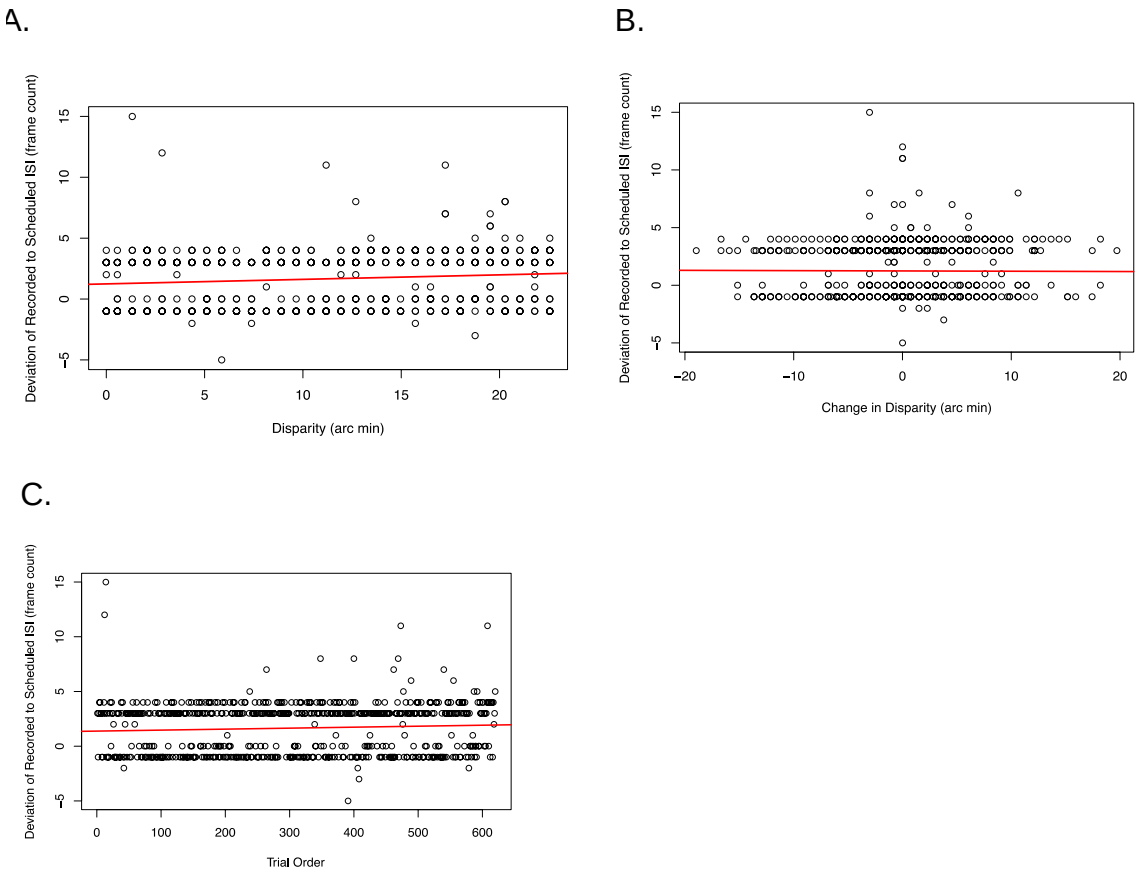

Figure 3. Recording 2: Plots of the difference between scheduled and recorded ISI values as a function of stimulus disparity (A) and as a function of the difference in disparity between two subsequent trials (B). Finally, (C) gives the ISI deviation over the sequence of trials in the recording (reflecting the progression of time). We note that the largest outliers occur on some of the earliest and later trials. All y-axes are in units of frame count. The red lines give the regression lines fit to the data.

5.3 Recording 3: Variable ISI, Variable Depth

The first two recordings suggest some positive correlation between the stimulus depth, or change in depth between trials, and the difference in scheduled and recorded ISI. The goal of Recording 3 is to combine the variable stimulus depth order with the variable scheduled ISI. This combination is consistent with the way the human observer studies are run: on each trial, the ISI is sampled from a random variable and then the depth of the stimulus is randomly selected from the set of possible values.

The order of disparity was replicated from the order used in Recording 2 in order to control for the variability within the change in depth between two subsequent trials. For each trial, the scheduled ISI duration was sampled from a uniform random variable ranging from 50–750 ms. All other aspects of the recording are the same as the earlier recordings. As Figure 4 illustrates, the ISI deviations were centered around zero with a concentration of data between a difference of -2 and 4 frames, consistent with the earlier recordings.

We again tested if the recorded ISI is greater than the scheduled ISI over the recording. Using a paired two-sample t-test, there was no significant difference on average between the recorded ISI ($M_{Recorded}$=381.8 ms) and the scheduled ISI ($M_{Scheduled}$=380.9 ms), t(1238)=0.08, p=.47.

A multiple regression analysis was used to test the predictive factors of disparity, change in disparity between trials, and the scheduled ISI on the ISI deviation. There were no significant predictors. Results are summarized in Table 1.

Finally, contradicting findings in Recording 2, a significant negative regression slope was found for ISI duration difference over as a function of trial order, which is a proxy for the total experiment duration (Figure 4; F(1, 618)=10.23, p<.05). Thus far we have found no practically significant effects. The change found across the duration of the experiment was a decrease in 0.002 camera frames as the experimental trials increased by one.

Table 1. Regression Analysis for Recording 3

Factor	B	t(615)	p-value
Disparity	-0.01	-0.264	.79
Change in Disparity	0.02	0.72	.47
Scheduled ISI	0.002	0.45	.66
Change in Scheduled ISI	-0.005	-1.53	.13

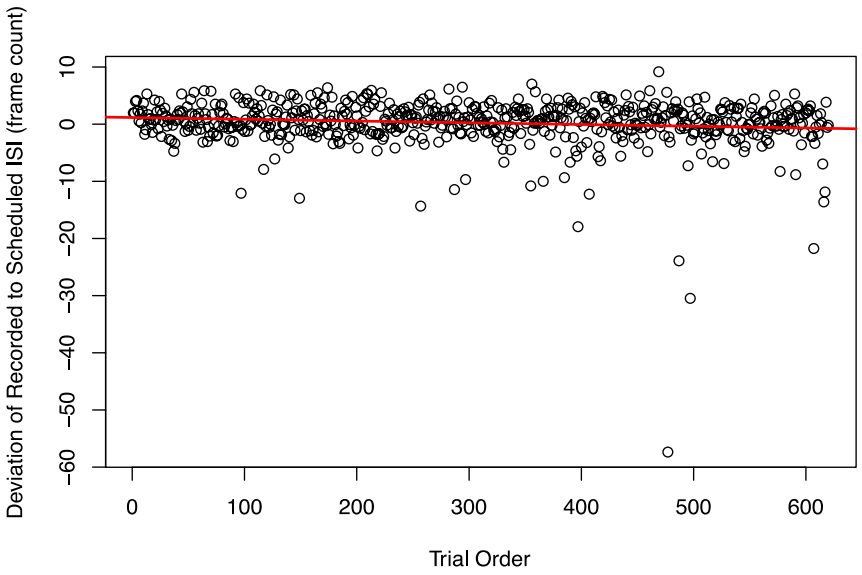

Figure 4. Recording 3: Plot of the ISI deviation as a function trial number in the experimental sequence. Note that the larger outliers in the data occur between the 450th and 515th trials. The red line is the regression line.

5.4 Recording 4: Machine Timing versus Display Timing

In Recordings 1-3, analysis focused on the display timing. However, the variability in the machine timing could be a source of variability in the display timing and the perceived ISI durations. The goal of Recording 4 was to record both the internal machine timing within the experimental code and the display timing for the stimuli, to identify if the machine timing biases the display timing. To do this, we added some additional controls to the sequence of stimuli, which is generated ahead of running the experiment. Usually this sequence is a random shuffle of the stimuli for each condition. In Recording 4, we controlled for the number disparity change magnitudes between subsequent trials, to have at least 40 samples from the range of possible values. The stimulus disparity cues were limited to 0, 5, 10, 15, and 20 arcmin. The possible change in disparity magnitudes were ±20, ±15, ±10, ±5, and ±0. This recording included a total of 1228 trials. The scheduled ISI was set to 50 ms.

The 3D stereo display can be converted to different types of the 3D displaying techniques (horizontal split type, vertical split type, etc.). Different 3D modes might have different display variability or interact differently with the machine timing. To remove this possible source of variance, we chose to not turn the 3D mode on and recorded the experiment using only one of the two images on the display in 2D mode.

Within our experimental code, a “tic-toc” timer was added. A tic time was recorded when the command for fixation offset executed, and the toc was recorded when the command for stimulus onset executed. The difference between toc and tic gives us the machine ISI time which we can compare to the recorded display ISI. This value should be equal to 50 ms if no additional machine variability is

occurring. Display ISI durations were computed by high-speed camera recording as in Recordings 1-3.

Two two-sample paired t-tests were conducted. In the first, we tested if the recorded ISI was greater than the machine (tic-toc) ISI against the null hypothesis that they are equal. There was a significant difference in the recorded ISI and the machine ISI ($M_{Machine}$=64.73 ms) which was greater than the recorded ISI, t(1227)=-20.68, p<.001.

In the second t-test, we examined if the recorded ISI was greater than the scheduled ISI. There was a significant difference between the recorded ISI ($M_{Recorded}$=56.4 ms) and the scheduled ISI ($M_{Scheduled}$=50.0 ms), t(1227)=23.74, p<.001. Thus, there were consistent differences in the average values of these various timing measurements. We now look for systematic variation in these differences.

A multiple regression model was fit to the ISI deviation between the machine tic-toc measurement and the recorded display measurement, in milliseconds. Overall, the model did not account for a significant proportion of the variability in the data, R^2=0.001, F(3,1224)=0.51, p=.68. The slopes for disparity and change in disparity between subsequent trials were not different from zero. ($B_{Disparity}$=-0.04, t(1225)=-0.64, p=.52; $B_{DisparityChange}$=0.06, t(1225)=1.21, p=.23).

Figure 5 shows the ISI deviation between the machine recording and the display recording as a function of trial number. It is clear that although there is a non-zero difference between these two measurements, the variability in that difference is small and that over time the difference is consistent. A regression line predicting the ISI deviation as a function of trial order showed a nearly zero slope for trial order, $B_{TrialOrder}$=-0.001, t(1226)=-1.21, p=.23. We infer that both the machine logging and display recording show a consistent baseline variation that is not a function of experimental factors.

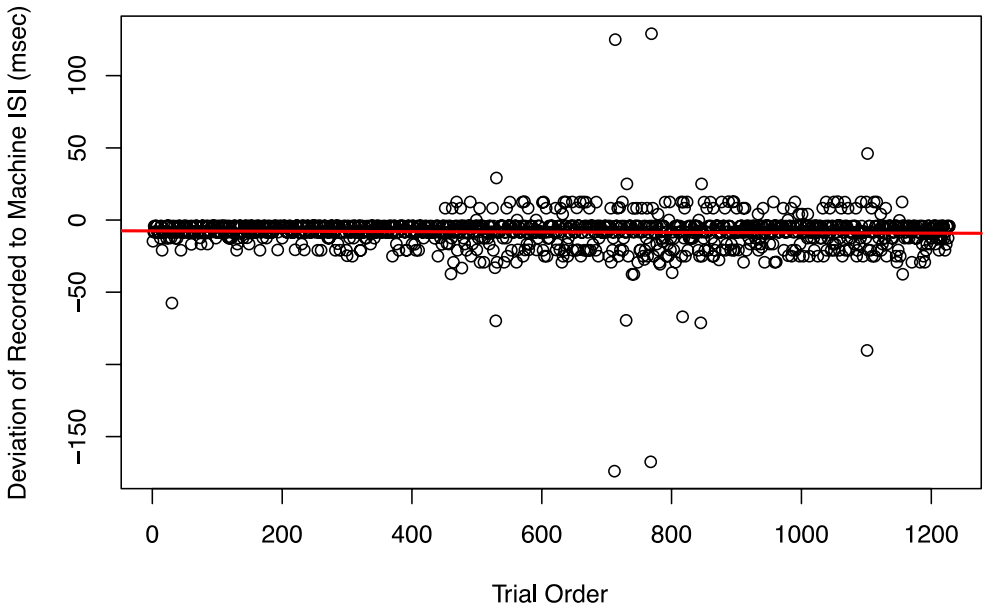

Figure 5. Recording 4: Plot of the difference between the machine ISI and the recorded display ISI as a function of the trial number. The red line is the regression line.

5.5 Recording 5: Screen Recording during Human Data Collection

Recordings 1-4 captured variability in the display ISI times inherent to the machine. To examine if this variability shows any predictive relationship with human performance, we recorded the display timing for a full experiment while a participant performed the stereo depth detection task. If there is no correlation between the human performance and the variability within the recorded experiment ISI, we can infer that the display variability is not systematically biasing the human observations. In this case, we can conclude that the display technology is sufficient for human stereo perception research.

The display recordings in Recording 5 were nearly identical to those in Recording 3, where we had the variation in stimulus depth and the randomly sampled ISI times. The camera position was adjusted to allow the participant to sit directly in front of the center of the screen. As discussed in the General Recording Methodology, the camera was positioned to the left of the participant, angled to the center of the display. The number of trials was 1200, equivalent to the human depth perception experiments of interest to 711 HPW/RHCV.

Two human observers were shown a series of circle-shaped objects on the LG 3D stereo display. One object appeared centrally on the display at a time. Their task was to determine whether or not each object was shown in depth (3D) or flat on the screen (2D). Participants were instructed to respond to every trial with either the left/right mouse button depending on if they thought each images was 2D or 3D. Right/left assignment was counterbalanced across participants. They were instructed to maintain a high level of accuracy, and while being accurate to respond as quickly as possible.

For Recording 5, data analyses were completed for each individual participant, not collapsed across participants. Note that some response time data for Participant 1 was lost after recording. Degrees of freedom in the subsequent analyses have been corrected for this.

Figure 6 shows the response times plotted against the ISI deviations for each participant. A correlation coefficient between these two variables was estimated for each participant. Participant 1's data had a correlation of r_{p1}=-0.067, and Participant 2 exhibiting a correlation of r_{p2}=.003.

Two t-tests were used to examine if the observed correlations between ISI deviation and response time were significantly different from a null hypothesis correlation of zero. No significant differences were found ($t_{p1}(879)$=01.98, p=0.48, $t_{p2}(1198)$=0.10, p=0.98). These results imply no practical relationship between the ISI deviations resulting from the display latency delays from the scheduled ISI, reported with recorded video using MATLAB, and the participant response times.

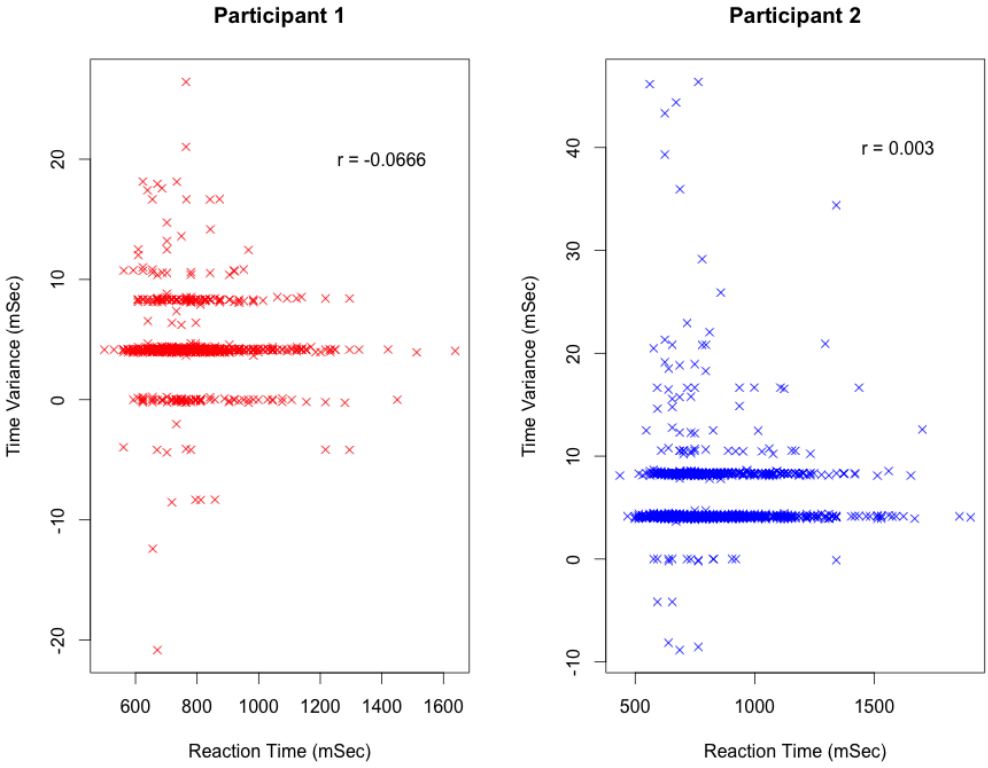

Figure 6. Recording 5: ISI deviation plotted against human response time on each trial. Data from Participant 1 is plotted on the left side, and data from Participant 2 is plotted on the right side.

Figure 7 shows the human response times plotted against the ISI frame count for all trials, together with the correlation between these two variables: r_{p1}=-0.017 and r_{p2}=0.027.

Two t-tests were used to test if the empirical correlation values were significantly different from zero. There were no significant differences between zero and the observed correlation values ($t_{p1}(879)$=-0.52, p=0.61, $t_{p2}(1198)$=-0.93, p=0.35). These results indicate no significant relationship between the recorded ISI and the response times of the participants.

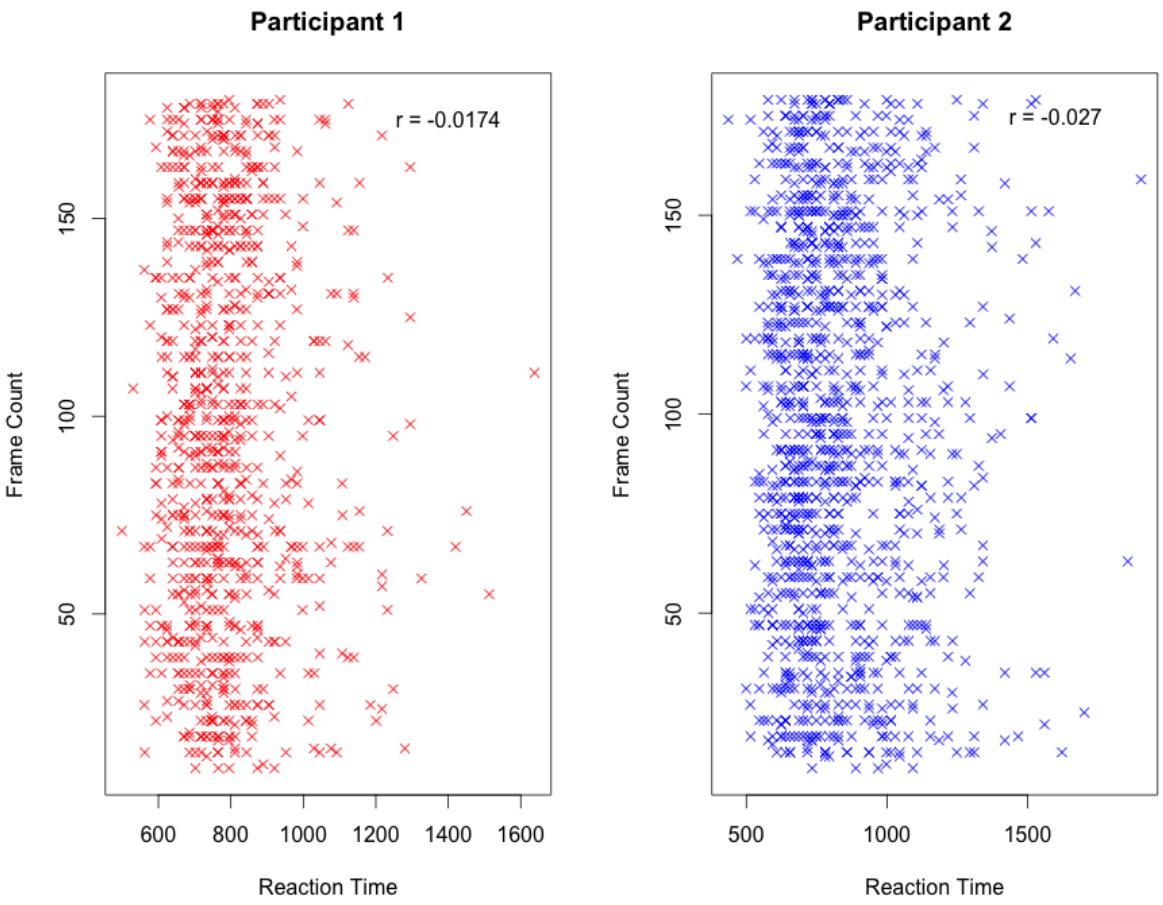

Figure 7. Recording 5: Response time plotted as a function of ISI frame count. Correlation coefficients are given in the upper right hand of each plot. Participant 1 is plotted on the left; Participant 2 is plotted on the left.

Figure 8 shows plots of the human observers' response times plotted against the stereo depth cue of the stimulus, for each trial. Correlation coefficients were estimated between these two variables: r_{p1}=-0.158 and r_{p2}=-0.381.

Two t-tests were used to test if the correlation coefficient estimates were significantly different from zero ($t_{p1}(879)$=-4.75, p<.001, $t_{p2}(1198)$=-14.26, p<.001). These tests indicate that the response times were significantly negatively correlated with stereo disparity. This means that responses were faster (shorter response times) when the stereo depth cues were larger. This is not surprising, as larger depth cues are easier to discriminate as 3D from the 2D stimuli.

Two additional paired two-sample t-tests were run to compare the ISI deviation as a function of stimulus stereo disparity. There was no significant difference in ISI deviation as a function of stimulus depth ($t_{1(}879)$=-.703, p=.482, $t_{2(}1198)$=-1.636, p=.102).

Taken together, there is not practical correlation between the human response times and the stimulus display timing. The effect of stimulus disparity on the human response times is consistent with the perception of stereo depth cues. Herein, we showed that this effect is not influenced by the display timing properties of those same images.

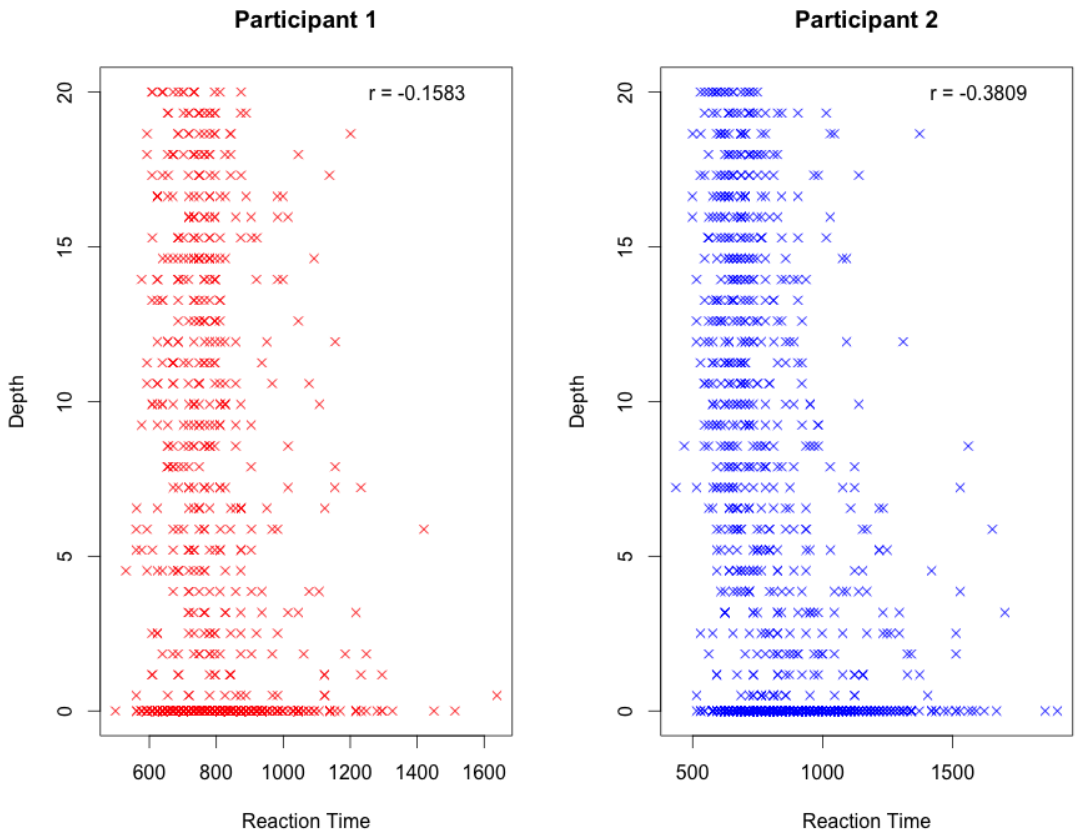

Figure 8. Recording 5: Response time plotted against of stimulus stereo disparity cue level. Correlation coefficients are given in the upper right of each plot. Participant 1 is plotted on the left side; Participant 2 is plotted on the right side.

6. Conclusion: Implications for 3D research

In conclusion, we were able to eliminate possible effects of specific ISI timings, trial-by-trial dependencies of disparity or ISI timing, and errors in software time recordings. We did find the 3D mode of the hardware display does cause an increase in ISI as disparity increases. However, these dependencies do not affect participant performance. The experiment under analysis can continue running due to no affects of participant performance.

To address the concerns some participants have brought to our attention, there may be a latency change in the perceptual processing of 3D images that gives the participants an experience of trials with larger depths taking an increased amount of time to display the image. There is very little literature addressing this concern of the timing of human stereoscopic processing. Some suggest that the verging of the eyes takes longer as the disparity increases (Marr, 1982; Wolf, Schuchardt, & Rosenzweig, 1996). This is a plausible explanation for the perceptual phenomenon. After finishing thorough hardware and software analyses, we can conclude the timing dependencies are not sufficient enough to be concerned for our experimental findings.

7. References

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: Freeman and Company.

R Core Team (2015). *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing.

Wolf, R., Schuchardt, M., & Rosenzweig, R. (1996). Depth illusion by delayed 3-D perception ('delayed stereopsis illusion'): A novel way to determine computation times in human vision by depth reversal in partially occluded moving objects. *Perception, 25, ECVP Abstract Supplement.*

LIST OF ACRONYMS, ABBREVIATIONS, AND SYMBOLS

2D	two-dimensional
3D	three-dimensional
711 HPW/RHCV	711 Human Performance Wing, Battlespace Visualization Branch
arcmin	Minute of arc, unit of angular measurement
fps	frame per second
ISI	Inter-stimulus Interval
ms	millisecond